Andres

#18

1646

RAW SEQUENCE LISTING DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:13

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT: Pepicelli, Carmen V
Lewis, Paula M
McMahon, Andrew P
(ii) TITLE OF INVENTION: Regulation of Lung Tissue by Hedgehog-like Polypeptides, and Formulations and Uses Related Thereto
(iii) NUMBER OF SEQUENCES: 24
(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Foley, Hoag & Eliot, LLP
(B) STREET: One Post Office Square
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02109
(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII(text)
C--> (vi) CURRENT APPLICATION DATA:
C--> (A) APPLICATION NUMBER: US/09/394,020A
C--> (B) FILING DATE: 10-Sep-1999
(C) CLASSIFICATION:
(vii) PRIOR APPLICATION DATA:
C--> (A) APPLICATION NUMBER: US 60/099,952
C--> (B) FILING DATE: 11-SEP-1998
(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Vincent, Matthew P
(B) REGISTRATION NUMBER: 36,709
(C) REFERENCE/DOCKET NUMBER: HUV-032.01
(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 832-1000
(B) TELEFAX: (617) 832-7000
(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1277 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1275
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
ATG GTC GAA ATG CTG CTG TTG ACA AGA ATT CTC TTG GTG GGC TTC ATC 48

RAW SEQUENCE LISTING DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:13

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

```
64 Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile
65  1               5                  10                  15
67 TGC GCT CTT TTA GTC TCC TCT GGG CTG ACT TGT GGA CCA GGC AGG GGC        96
68 Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly
69              20                  25                  30
71 ATT GGA AAA AGG AGG CAC CCC AAA AAG CTG ACC CCG TTA GCC TAT AAG       144
72 Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
73          35                  40                  45
75 CAG TTT ATT CCC AAT GTG GCA GAG AAG ACC CTA GGG GCC AGT GGA AGA       192
76 Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg
77      50                  55                  60
79 TAT GAA GGG AAG ATC ACA AGA AAC TCC GAG AGA TTT AAA GAA CTA ACC       240
80 Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr
81  65                  70                  75                  80
83 CCA AAT TAC AAC CCT GAC ATT ATT TTT AAG GAT GAA GAG AAC ACG GGA       288
84 Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly
85                  85                  90                  95
87 GCT GAC AGA CTG ATG ACT CAG CGC TGC AAG GAC AAG CTG AAT GCC CTG       336
88 Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu
89              100                 105                 110
91 GCG ATC TCG GTG ATG AAC CAG TGG CCC GGG GTG AAG CTG CGG GTG ACC       384
92 Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr
93          115                 120                 125
95 GAG GGC TGG GAC GAG GAT GGC CAT CAC TCC GAG GAA TCG CTG CAC TAC       432
96 Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr
97      130                 135                 140
99 GAG GGT CGC GCC GTG GAC ATC ACC ACG TCG GAT CGG GAC CGC AGC AAG       480
100 Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys
101 145                 150                 155                 160
103 TAC GGA ATG CTG GCC CGC CTC GCC GTC GAG GCC GGC TTC GAC TGG GTC       528
104 Tyr Gly Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val
105                 165                 170                 175
107 TAC TAC GAG TCC AAG GCG CAC ATC CAC TGC TCC GTC AAA GCA GAA AAC       576
108 Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn
109             180                 185                 190
111 TCA GTG GCA GCG AAA TCA GGA GGC TGC TTC CCT GGC TCA GCC ACA GTG       624
112 Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val
113         195                 200                 205
115 CAC CTG GAG CAT GGA GGC ACC AAG CTG GTG AAG GAC CTG AGC CCT GGG       672
116 His Leu Glu His Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly
117     210                 215                 220
119 GAC CGC GTG CTG GCT GCT GAC GCG GAC GGC CGG CTG CTC TAC AGT GAC       720
120 Asp Arg Val Leu Ala Ala Asp Ala Asp Gly Arg Leu Leu Tyr Ser Asp
121 225                 230                 235                 240
123 TTC CTC ACC TTC CTC GAC CGG ATG GAC AGC TCC CGA AAG CTC TTC TAC       768
124 Phe Leu Thr Phe Leu Asp Arg Met Asp Ser Ser Arg Lys Leu Phe Tyr
125                 245                 250                 255
127 GTC ATC GAG ACG CGG CAG CCC CGG GCC CGG CTG CTA CTG ACG GCG GCC       816
128 Val Ile Glu Thr Arg Gln Pro Arg Ala Arg Leu Leu Leu Thr Ala Ala
```

RAW SEQUENCE LISTING DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:14

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

```
129             260                 265                 270
131 CAC CTG CTC TTT GTG GCC CCC CAG CAC AAC CAG TCG GAG GCC ACA GGG       864
132 His Leu Leu Phe Val Ala Pro Gln His Asn Gln Ser Glu Ala Thr Gly
133         275                 280                 285
135 TCC ACC AGT GGC CAG GCG CTC TTC GCC AGC AAC GTG AAG CCT GGC CAA       912
136 Ser Thr Ser Gly Gln Ala Leu Phe Ala Ser Asn Val Lys Pro Gly Gln
137     290                 295                 300
139 CGT GTC TAT GTG CTG GGC GAG GGC GGG CAG CAG CTG CTG CCG GCG TCT       960
140 Arg Val Tyr Val Leu Gly Glu Gly Gly Gln Gln Leu Leu Pro Ala Ser
141 305                 310                 315                 320
143 GTC CAC AGC GTC TCA TTG CGG GAG GAG GCG TCC GGA GCC TAC GCC CCA      1008
144 Val His Ser Val Ser Leu Arg Glu Glu Ala Ser Gly Ala Tyr Ala Pro
145                 325                 330                 335
147 CTC ACC GCC CAG GGC ACC ATC CTC ATC AAC CGG GTG TTG GCC TCC TGC      1056
148 Leu Thr Ala Gln Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys
149             340                 345                 350
151 TAC GCC GTC ATC GAG GAG CAC AGT TGG GCC CAT TGG GCC TTC GCA CCA      1104
152 Tyr Ala Val Ile Glu Glu His Ser Trp Ala His Trp Ala Phe Ala Pro
153         355                 360                 365
155 TTC CGC TTG GCT CAG GGG CTG CTG GCC GCC CTC TGC CCA GAT GGG GCC      1152
156 Phe Arg Leu Ala Gln Gly Leu Leu Ala Ala Leu Cys Pro Asp Gly Ala
157     370                 375                 380
159 ATC CCT ACT GCC GCC ACC ACC ACC ACT GGC ATC CAT TGG TAC TCA CGG      1200
160 Ile Pro Thr Ala Ala Thr Thr Thr Thr Gly Ile His Trp Tyr Ser Arg
161 385                 390                 395                 400
162 CTC CTC TAC CGC ATC GGC AGC TGG GTG CTG GAT GGT GAC GCG CTG CAT      1248
163 Leu Leu Tyr Arg Ile Gly Ser Trp Val Leu Asp Gly Asp Ala Leu His
164                 405                 410                 415
166 CCG CTG GGC ATG GTG GCA CCG GCC AGC TG                              1277
167 Pro Leu Gly Met Val Ala Pro Ala Ser
168             420                 425
171 (2) INFORMATION FOR SEQ ID NO: 2:
173      (i) SEQUENCE CHARACTERISTICS:
174           (A) LENGTH: 1190 base pairs
175           (B) TYPE: nucleic acid
176           (C) STRANDEDNESS: both
177           (D) TOPOLOGY: linear
179     (ii) MOLECULE TYPE: cDNA
181     (ix) FEATURE:
182           (A) NAME/KEY: CDS
183           (B) LOCATION: 1..1191
185     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
187 ATG GCT CTG CCG GCC AGT CTG TTG CCC CTG TGC TGC TTG GCA CTC TTG        48
188 Met Ala Leu Pro Ala Ser Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu
189   1               5                  10                  15
191 GCA CTA TCT GCC CAG AGC TGC GGG CCG GGC CGA GGA CCG GTT GGC CGG        96
192 Ala Leu Ser Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg
193              20                  25                  30
195 CGG CGT TAT GTG CGC AAG CAA CTT GTG CCT CTG CTA TAC AAG CAG TTT       144
```

RAW SEQUENCE LISTING DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:14

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

```
196 Arg Arg Tyr Val Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe
197         35                  40                  45
199 GTG CCC AGT ATG CCC GAG CGG ACC CTG GGC GCG AGT GGG CCA GCG GAG       192
200 Val Pro Ser Met Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu
201      50                  55                  60
203 GGG AGG GTA ACA AGG GGG TCG GAG CGC TTC CGG GAC CTC GTA CCC AAC       240
204 Gly Arg Val Thr Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn
205  65                  70                  75                  80
207 TAC AAC CCC GAC ATA ATC TTC AAG GAT GAG GAG AAC AGC GGC GCA GAC       288
208 Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp
209                 85                  90                  95
211 CGC CTG ATG ACA GAG CGT TGC AAA GAG CGG GTG AAC GCT CTA GCC ATC       336
212 Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile
213             100                 105                 110
215 GCG GTG ATG AAC ATG TGG CCC GGA GTA CGC CTA CGT GTG ACT GAA GGC       384
216 Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
217         115                 120                 125
219 TGG GAC GAG GAC GGC CAC CAC GCA CAG GAT TCA CTC CAC TAC GAA GGC       432
220 Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly
221     130                 135                 140
223 CGT GCC TTG GAC ATC ACC ACG TCT GAC CGT GAC CGT AAT AAG TAT GGT       480
224 Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly
225 145                 150                 155                 160
227 TTG TTG GCG CGC CTA GCT GTG GAA GCC GGA TTC GAC TGG GTC TAC TAC       528
228 Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
229                 165                 170                 175
231 GAG TCC CGC AAC CAC ATC CAC GTA TCG GTC AAA GCT GAT AAC TCA CTG       576
232 Glu Ser Arg Asn His Ile His Val Ser Val Lys Ala Asp Asn Ser Leu
233             180                 185                 190
235 GCG GTC CGA GCC GGA GGC TGC TTT CCG GGA AAT GCC ACG GTG CGC TTG       624
236 Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu
237         195                 200                 205
239 CGG AGC GGC GAA CGG AAG GGG CTG AGG GAA CTA CAT CGT GGT GAC TGG       672
240 Arg Ser Gly Glu Arg Lys Gly Leu Arg Glu Leu His Arg Gly Asp Trp
241     210                 215                 220
243 GTA CTG GCC GCT GAT GCA GCG GGC CGA GTG GTA CCC ACG CCA GTG CTG       720
244 Val Leu Ala Ala Asp Ala Ala Gly Arg Val Val Pro Thr Pro Val Leu
245 225                 230                 235                 240
247 CTC TTC CTG GAC CGG GAT CTG CAG CGC CGC GCC TCG TTC GTG GCT GTG       768
248 Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val
249                 245                 250                 255
251 GAG ACC GAG CGG CCT CCG CGC AAA CTG TTG CTC ACA CCC TGG CAT CTG       816
252 Glu Thr Glu Arg Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu
253             260                 265                 270
255 GTG TTC GCT GCT CGC GGG CCA GCG CCT GCT CCA GGT GAC TTT GCA CCG       864
256 Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro
257         275                 280                 285
259 GTG TTC GCG CGC CGC TTA CGT GCT GGC GAC TCG GTG CTG GCT CCC GGC       912
260 Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly
```

RAW SEQUENCE LISTING DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:14

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

```
261     290                 295                 300
263 GGG GAC GCG CTC CAG CCG GCG CGC GTA GCC CGC GTG GCG CGC GAG GAA       960
264 Gly Asp Ala Leu Gln Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu
265 305                 310                 315                 320
267 GCC GTG GGC GTG TTC GCA CCG CTC ACT GCG CAC GGG ACG CTG CTG GTC      1008
268 Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
269                 325                 330                 335
271 AAC GAC GTC CTC GCC TCC TGC TAC GCG GTT CTA GAG AGT CAC CAG TGG      1056
272 Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
273             340                 345                 350
275 GCC CAC CGC GCC TTC GCC CCT TTG CGG CTG CTG CAC GCG CTC GGG GCT      1104
276 Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala
277         355                 360                 365
279 CTG CTC CCT GGG GGT GCA GTC CAG CCG ACT GGC ATG CAT TGG TAC TCT      1152
280 Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser
281     370                 375                 380
283 CGC CTC CTT TAC CGC TTG GCC GAG GAG TTA ATG GGC TG                   1190
284 Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Met Gly
285 385                 390                 395
288 (2) INFORMATION FOR SEQ ID NO: 3:
290      (i) SEQUENCE CHARACTERISTICS:
291           (A) LENGTH: 1281 base pairs
292           (B) TYPE: nucleic acid
293           (C) STRANDEDNESS: both
294           (D) TOPOLOGY: linear
296     (ii) MOLECULE TYPE: cDNA
299     (ix) FEATURE:
300           (A) NAME/KEY: CDS
301           (B) LOCATION: 1..1233
303     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
305 ATG TCT CCC GCC TGG CTC CGG CCC CGA CTG CGG TTC TGT CTG TTC CTG        48
306 Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu
307   1                 5                  10                  15
309 CTG CTG CTG CTT CTG GTG CCG GCG GCG CGG GGC TGC GGG CCG GGC CGG        96
310 Leu Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg
311             20                  25                  30
313 GTG GTG GGC AGC CGC CGG AGG CCG CCT CGC AAG CTC GTG CCT CTT GCC       144
314 Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala
315         35                  40                  45
317 TAC AAG CAG TTC AGC CCC AAC GTG CCG GAG AAG ACC CTG GGC GCC AGC       192
318 Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser
319       50                  55                  60
321 GGG CGC TAC GAA GGC AAG ATC GCG CGC AGC TCT GAG CGC TTC AAA GAG       240
322 Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu
323  65                  70                  75                  80
325 CTC ACC CCC AAC TAC AAT CCC GAC ATC ATC TTC AAG GAC GAG GAG AAC       288
326 Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn
327                 85                  90                  95
329 ACG GGT GCC GAC CGC CTC ATG ACC CAG CGC TGC AAG GAC CGT CTG AAC       336
```

VERIFICATION SUMMARY DATE: 07/18/2000
PATENT APPLICATION: US/09/394,020A TIME: 13:52:15

Input Set : A:\SeqLstg.txt
Output Set: N:\CRF3\07182000\I394020A.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(vi) CURRENT APPLICATION DATA:]
L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:36 M:220 C: Keyword misspelled or invalid format, (A) APPLICATION NUMBER:
L:37 M:220 C: Keyword misspelled or invalid format, (B) FILING DATE:
L:801 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:1759 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:2287 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2293 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2305 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2311 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:2346 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2349 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2358 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2367 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2376 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:2385 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=23, Value=[not relevant]